	1				50
EG327	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAS
BZ198	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
BZ10	MNKISRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
H15	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
EG329	MNEILRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAS
PMC21	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAS
H38	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
P20	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLSATVQAN
Z2491	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
H41	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
Consensus	MN-I-RIIWN	SALNAWV-VS	ELTRNHTKRA	SATV-TAVLA	TLL-ATVQA-

C1

	51				100
EG327	TTDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.VTED	SNWGVYFDKK
BZ198	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
BZ10	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
H15	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
EG329	ANNEEQEEDL	YLDPVLRTVA	VLIVNSDKEG	TGEKEKVEEN	SDWAVYFNEK
PMC21	ANNEEQEEDL	YLDPVQRTVA	VLIVNSDKEG	TGEKEKVEEN	SDWAVYFNEK
H38	ATDED..EEE	ELEPVVRSAL	VLQFMIDKEG	NGENE.STGN	IGWSIYYDNH
P20	ATDTD..EDE	ELESVARSAL	VLQFMIDKEG	NGEIESTGDI	GWSIYYDDHN
Z2491	ATDED..EEE	ELESVQR.SV	VGSIQASMEG	SGELET...I	SLSMTNDSKE
H41	ATDED..EEE	ELESVQR.SV	VGSIQASMEG	SVELET...I	SLSMTNDSKE
Consensus	----------	-L--V-R---	V-------EG	--E-E-----	----------

V1

	101				150
EG327	GVLTAGTITL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTDLTSVG
BZ198	RVLKAGAITL	KAGDNLKIKQ	NTNENTNDSS	....FTYSLK	KDLTDLTSVE
BZ10	RVLKAGAITL	KAGDNLKIKQ	NTNENTNENT	NDSSFTYSLK	KDLTDLTSVE
H15	RVLKAGAITL	KAGDNLKIKQ	NTNENTNENT	NDSSFTYSLK	KDLTDLTSVE
EG329	GVLTAREITL	KAGDNLKIKQ	NG...TN...	....FTYSLK	KDLTDLTSVG
PMC21	GVLTAREITL	KAGDNLKIKQ	NG...TN...	....FTYSLK	KDLTDLTSVG
H38	NTLHGATVTL	KAGDNLKIKQ	NTNKNTNENT	NDSSFTYSLK	KDLTDLTSVE
P20	TLHG.ATVTL	KAGDNLKIKQ	SGKD......	....FTYSLK	KELKDLTSVE
Z2491	FVDPYIVVTL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTGLINVE
H41	FVDPYIVVTL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTGLINVE
Consensus	--------TL	KAGDNLKIKQ	----------	----FTYSLK	K-L--L--V-

V1 C2 V2 C3

FIG. 1A

```
          151                                                       200
EG327     TEKLSFSANS NKVNITSDTK GLNFAKKTAE TNGDTTVHLN GIGSTLTDTL
BZ198     TEKLSFGANG NKVNITSDTK GLNFAKETAG TNGDPTVHLN GIGSTLTDTL
BZ10      TEKLSFGANG NKVNITSDTK GLNFAKETAG TNGDPTVHLN GIGSTLTDTL
H15       TEKLSFGANG NKVNITSDTK GLNFAKETAG TNGDPTVHLN GIGSTLTDTL
EG329     TEKLSFSANG NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL
PMC21     TEKLSFSANG NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL
H38       TEKLSFGANG NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL
P20       TEKLSFGANG NKVNITSDTK GLNFAKETAG TNGDPTVHLN GIGSTLTDTL
Z2491     TEKLSFGANG KKVNIISDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL
H41       TEKLSFGANG KKVNIISDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDML
Consensus TEKLSF-AN- -KVNI-SDTK GLNFAK-TA- TNGD-TVHLN GIGSTLTD-L
                                    C3

          201                                                       250
EG327     LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
BZ198     LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
BZ10      LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
H15       LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
EG329     LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
PMC21     LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
H38       LNTGATTNVT NDNVTDDKKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
P20       AGSSASHVDA GNQST..HYT RAASIKDVLN AGWNIKGVKT GSTTGQSENV
Z2491     AGSSASHVDA GNQST..HYT RAASIKDVLN AGWNIKGVKT GSTTGQSENV
H41       LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
Consensus ----A----- -----T----- RAAS-KDVLN AGWNIKGVK- G-T-----NV
               V3                          C4                V4

          251                                                       300
EG327     DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
BZ198     DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
BZ10      DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
H15       DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
EG329     DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
PMC21     DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
H38       DFVHTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
P20       DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
Z2491     DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
H41       DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
Consensus DFV-TYDTVE FLSADTKTTT VNVESKDNGK -TEVKIGAKT SVIKEKDGKL
                                    C5
```

FIG. 1B

	301				350
EG327	VTGKDKGEND	SSTDKGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
BZ198	VTGKGKDENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
BZ10	VTGKGKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
H15	VTGKGKDENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
EG329	VTGKDKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
PMC21	VTGKDKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
H38	VTGKGKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
P20	VTGKGKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
Z2491	VTGKGKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
H41	VTGKGKGENG	SSTDEGEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK
Consensus	VTGK-K-EN-	SSTD-GEGLV	TAKEVIDAVN	KAGWRMKTTT	ANGQTGQADK

C5

	351				400
EG327	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVMYDVNVGD	ALNVNQLQNS
BZ198	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
BZ10	FETVTSGTKV	TFASGNGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
H15	FETVTSGTKV	TFASGNGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
EG329	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVMYDVNVGD	ALNVNQLQNS
PMC21	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVMYDVNVGD	ALNVNQLQNS
H38	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
P20	FETVTSGTKV	TFASGNGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
Z2491	FETVTSGTNV	TFASGKGTTA	TVSKDDQGNI	TVMYDVNVGD	ALNVNQLQNS
H41	FETVTSGTKV	TFASGNGTTA	TVSKDDQGNI	TVKYDVNVGD	ALNVNQLQNS
Consensus	FETVTSGT-V	TFASG-GTTA	TVSKDDQGNI	TV-YDVNVGD	ALNVNQLQNS

C5

	401				450
EG327	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
BZ198	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
BZ10	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
H15	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
EG329	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
PMC21	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
H38	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
P20	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
Z2491	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EISRNGKNID
H41	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EITRNGKNID
Consensus	GWNLDSKAVA	GSSGKVISGN	VSPSKGKMDE	TVNINAGNNI	EI-RNGKNID

C5

FIG. 1C

```
          451                                                  500
EG327     IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
BZ198     IATSMAPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDTNK PVRITNVAPG
BZ10      IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
H15       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
EG329     IATSMTPQFS SVSLGAGADA PTLSVDG.DA LNVGSKKDNK PVRITNVAPG
PMC21     IATSMTPQFS SVSLGAGADA PTLSVDG.DA LNVGSKKDNK PVRITNVAPG
H38       IATSMTPQFS SVSLGAGADA PTLSVDDKGA LNVGSKDANK PVRITNVAPG
P20       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
Z2491     IATSMAPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
H41       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
Consensus IATSM-PQFS SVSLGAGADA PTLSVD---A LNVGSK--NK PVRITNVAPG
                                                      C5

          501                                                  550
EG327     VKEGDVTNVA QLKGVAQNLN NHIDNVDGNA RAGIAQAIAT AGLVQAYLPG
BZ198     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
BZ10      VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLAQAYLPG
H15       VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLAQAYLPG
EG329     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
PMC21     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
H38       VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
P20       VKEGDVTNVA QLKGVAQNLN NRIDNVNGNA RAGIAQAIAT AGLAQAYLPG
Z2491     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
H41       VKEGDVTNVA QLKGVAQNLN NRIDNVNGNA RAGIAQAIAT AGLVQAYLPG
Consensus VKEGDVTNVA QLKGVAQNLN N-IDNV-GNA RAGIAQAIAT AGL-QAYLPG
                                                      C5

          551                                                  600
EG327     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
BZ198     KSMMAIGGDT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
BZ10      KSMMAIGGGT YRGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGTSASV
H15       KSMMAIGGGT YRGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGASASV
EG329     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
PMC21     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
H38       KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
P20       KSMMAIGGGT YLGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGTSASV
Z2491     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
H41       KSMMAIGGGT YLGEAGYAIG YSSISAGGNW IIKGTASGNS RGHFGASASV
Consensus KSMMAIGG-T Y-GEAGYAIG YSSIS--GNW -IKGTASGNS RGHFG-SASV
                                                      C5
```

FIG. 1D

	601
EG327	GYQW.
BZ198	GYQW.
BZ10	GYQW.
H15	GYQW.
EG329	GYQW.
PMC21	GYQW.
H38	GYQW.
P20	GYQW.
Z2491	GYQW.
H41	GYQW.
Consensus	GYQW.
	C5

FIG. 1E

```
           1                                                                    70
H15        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
BZ10       ATGAACAAAA TATCCCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
BZ198      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
P20        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT AGTCGTATCC GAGCTCACAC
H38        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
Z2491      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
H41        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
EG329      ATGAACGAAA TATTGCGCAT CATTTGGAAT AGCGCCCTCA ATGCCTGGGT CGTTGTATCC GAGCTCACAC
PMC21      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT CGTCGTATCC GAGCTCACAC
EG327      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
Consensus  ATGAAC-AAA TAT--CGCAT CATTTGGAAT AG-GCCCTCA ATGC-TGGGT -G--GTATCC GAGCTCACAC
                                           C1

           71                                                                   140
H15        GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
BZ10       GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
BZ198      GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
P20        GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGCTGT CCGCAACGGT
H38        GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACGCTGTTGT TTGCAACGGT
Z2491      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
H41        GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
EG329      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT
PMC21      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT
EG327      GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
Consensus  GCAACCACAC CAAACGCGCC TCCGCAACCG TG--GACCGC CGTATTGGCG AC-CTG-TGT --GCAACGGT
                                           C1

           141                                                                  210
H15        TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
BZ10       TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
BZ198      TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
P20        TCAGGCGAAT GCTACCGATA CCGAT..... .GAAGATGAA GAGTTAGAAT CCGTAGCACG CTCTGCTCTG
H38        TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAC CCGTAGTACG CTCTGCTCTG
Z2491      TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAT CCGTACAACG CTCTGTCGTA
H41        TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAT CCGTACAACG CTCTG...TC
EG329      TCAGGCAAGT GCTAACAATG AAGAGCAAGA AGAAGATTTA TATTTAGACC CCGTGCTACG CACTGTTGCC
PMC21      TCAGGCAAGT GCTAACAATG AAGAGCAAGA AGAAGATTTA TATTTAGACC CCGTACAACG CACTGTTGCC
EG327      TCAGGCGAGT ACTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
Consensus  TCAGGC-A-T -CTA-C-AT- --GA------ ----GA---A -A-TTAGA-- CCGT---ACG C-CTG-----
           C1                                   V1
```

FIG. 2A

```
           211                                                                       280
H15        GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
BZ10       GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
BZ198      GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
P20        GTGTTGCAAT TCATGATCGA TAAAGAAGGC AATGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
H38        GTGTTGCAAT TCATGATCGA TAAAGAAGGC AATGGAGAAA TCGAATCTAC AGGAGA...T ATAGGTTGGA
Z2491      GGG..AGCAT TCAAG.CCAG TATGGAAGGC AGCGGCGAAT TGGAAACGAT AGGAAA...T ATAGGTTGGA
H41        GTAGGGAGCA TTCAAGCCAG TATGGAAGGC AGCGTCGAAT TGGAAACGAT ATCAT....T ATCAATGACT
EG329      GTGTTGATAG TCAATTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAAAAGT A......... TCATTATCAA
PMC21      GTGTTGATAG TCAATTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAAAAGT AGAAGAAAAT TCAGATTGGG
EG327      GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGTTAC AGAAGA...T TCAAATTGGG
Consensus  G--------- T------C-- TA--GAAGGC A--G--GAA- --GAA----- A--------- ----------
                                        V1

           281                                                                       350
H15        CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
BZ10       CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
BZ198      CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
P20        GTATATATTA CGACGATCAC AACACTCTAC ACGGCGCAAC CGTTACCCTC AAAGCCGGCG ACAACCTGAA
H38        GTATATATTA CGACAATCAC AACACTCTAC ACGGCGCAAC CGTTACCCTC AAAGCCGGCG ACAACCTGAA
Z2491      AACGACAGCA AGGAATTTGT AGACCCATAC ATAGTA.... .GTTACCCTC AAAGCCGGCG ACAACCTGAA
H41        TGACTAACGA CAGCAAGGAA TTTGTAGACC CATACATAGT AGTTACCCTC AAAGCCGGCG ACAACCTGAA
EG329      CAGTATATTT CAACGAGAAA GGAGTACTAA CAGCCAGAGA AATCACCCTC AAAGCCGGCG ACAACCTGAA
PMC21      CAGTATATTT CAACGAGAAA GGAGTACTAA CAGCCAGAGA AATCACCCTC AAAGCCGGCG ACAACCTGAA
EG327      GAGTATATTT CGACAAGAAA GGAGTACTAA CAGCCGGAAC AATCACCCTC AAAGCCGGCG ACAACCTGAA
Consensus  ------A--- ---------- ---------- ---------- --T-ACCCTC AAAGCCGGCG ACAACCTGAA
                                        V1                        C2

           351                                                                       420
H15        AATCAAACAA AACACCAATG AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCCCTGAAA
BZ10       AATCAAACAA AACACCAATG AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCCCTGAAA
BZ198      AATCAAACAA AACACCAATG AAAACACC.. .......... AATGACAGTA GCTTCACCTA CTCCCTGAAA
P20        AATCAAACAA AGCGGCAAAG A......... .......... .......... .CTTCACCTA CTCGCTGAAA
H38        AATCAAACAA AACACCAATA AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCGCTGAAA
Z2491      AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
H41        AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
EG329      AATCAAACAA AAC....... .......G.. .......... ....GCACAA ACTTCACCTA CTCGCTGAAA
PMC21      AATCAAACAA AAC....... .......G.. .......... ....GCACAA ACTTCACCTA CTCGCTGAAA
EG327      AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
Consensus  AATCAAACAA A-C------- ---------- ---------- ---------- -CTTCACCTA CTC-CTGAAA
             C2                 V2                                               C3
```

FIG. 2B

	421						490
H15	AAAGACCTCA	CAGATCTGAC	CAGTGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGT	AATAAAGTCA
BZ10	AAAGACCTCA	CAGATCTGAC	CAGTGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGT	AATAAAGTCA
BZ198	AAAGACCTCA	CAGATCTGAC	CAGTGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGT	AATAAAGTCA
P20	AAAGAGCTGA	AAGACCTGAC	CAGTGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGT	AATAAAGTCA
H38	AAAGACCTCA	CAGATCTGAC	CAGTGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGC	AATAAAGTCA
Z2491	AAAGACCTCA	CAGGCCTGAT	CAATGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGC	AAGAAAGTCA
H41	AAAGACCTCA	CAGGCCTGAT	CAATGTTGAA	ACTGAAAAAT	TATCGTTTGG	CGCAAACGGC	AAGAAAGTCA
EG329	AAAGACCTCA	CAGATCTGAC	CAGTGTTGGA	ACTGAAAAAT	TATCGTTTAG	CGCAAACGGC	AATAAAGTCA
PMC21	AAAGACCTCA	CAGATCTGAC	CAGTGTTGGA	ACTGAAAAAT	TATCGTTTAG	CGCAAACGGC	AATAAAGTCA
EG327	AAAGACCTCA	CAGATCTGAC	CAGTGTTGGA	ACTGAAAAAT	TATCGTTTAG	CGCAAACAGC	AATAAAGTCA
Consensus	AAAGA-CT-A	-AG--CTGA-	CA-TGTTG-A	ACTGAAAAAT	TATCGTTT-G	CGCAAAC-G-	AA-AAAGTCA

C3

	491						560
H15	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACCCCACGGT
BZ10	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACCCCACGGT
BZ198	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACCCCACGGT
P20	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACCCCACGGT
H38	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TCGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACACCACGGT
Z2491	ACATCATAAG	CGACACCAAA	GGCTTGAATT	TCGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACACCACGGT
H41	ACATCATAAG	CGACACCAAA	GGCTTGAATT	TCGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACACCACGGT
EG329	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACACCACGGT
PMC21	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TTGCGAAAGA	AACGGCTGGG	ACGAACGGCG	ACACCACGGT
EG327	ACATCACAAG	CGACACCAAA	GGCTTGAATT	TCGCGAAAAA	AACGGCTGAG	ACCAACGGCG	ACACCACGGT
Consensus	ACATCA-AAG	CGACACCAAA	GGCTTGAATT	T-GCGAAA-A	AACGGCTG-G	AC-AACGGCG	AC-CCACGGT

C3

	561						630
H15	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
BZ10	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
BZ198	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
P20	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTT	GCGGGTTCTT	CTGCTTCTCA	CGTTGATGCG
H38	TCATCTGAAC	GGTATTGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
Z2491	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTT	GCGGGTTCTT	CTGCTTCTCA	CGTTGATGCG
H41	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATATGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
EG329	TCATCTGAAC	GGTATTGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
PMC21	TCATCTGAAC	GGTATTGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
EG327	TCATCTGAAC	GGTATCGGTT	CGACTTTGAC	CGATACGCTG	CTGAATACCG	GAGCGACCAC	AAACGTAACC
Consensus	TCATCTGAAC	GGTAT-GGTT	CGACTTTGAC	CGATA-GCT-	--G--T-C--	--GC--C---	----G---C-

C3 V3

FIG. 2C

```
           631                                                                        700
H15        AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
BZ10       AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
BZ198      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
P20        GGTAACCAAA GTACACATTA C......ACT CGTGCAGCAA GTATTAAGGA TGTGTTGAAT GCGGGTTGGA
H38        AACGACAACG TTACCGATGA CAAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
Z2491      GGTAACCAAA GTACACATTA C......ACT CGTGCAGCAA GTATTAAGGA TGTGTTGAAT GCGGGTTGGA
H41        AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
EG329      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCTGGCTGGA
PMC21      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCTGGCTGGA
EG327      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
Consensus  ----AC-A-- -TAC--AT-A C------A-- CGTGC-GCAA G--TTAA-GA -GT-TT-AA- GC-GG-TGGA
                 V3                                             C4

           701                                                                        770
H15        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
BZ10       ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTC GATTTCGTCC GCACTTACGA
BZ198      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
P20        ATATTAAGGG TGTTAAAACT GGCTCAACAA CTGGTCAATC AGAAAATGTC GATTTCGTCC GCACTTACGA
H38        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC ACACTTACGA
Z2491      ATATTAAGGG TGTTAAAACT GGCTCAACAA CTGGTCAATC AGAAAATGTC GATTTCGTCC GCACTTACGA
H41        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
EG329      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
PMC21      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
EG327      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
Consensus  A-ATTAA-GG -GTTAAA-C- GG--CAACA- CT------TC -GA-AA-GT- GATTTCGTCC -CACTTACGA
                C4                V4                                  C5

           771                                                                        840
H15        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
BZ10       CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
BZ198      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
P20        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
H38        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
Z2491      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
H41        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
EG329      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
PMC21      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
EG327      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
Consensus  CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
                                                  C5

           841                                                                        910
H15        AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
BZ10       AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
BZ198      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
P20        AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
H38        AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
Z2491      AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
H41        AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
EG329      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
PMC21      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
EG327      AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATCA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
Consensus  A-AACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTAT-A AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
                                                  C5
```

FIG. 2D

	911						980
H15	AAGGCAAAGA	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
BZ10	AAGGCAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
BZ198	AAGGCAAAGA	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
P20	AAGGCAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
H38	AAGGCAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
Z2491	AAGGCAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
H41	AAGGCAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
EG329	AAGACAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
PMC21	AAGACAAAGG	CGAGAATGGT	TCTTCTACAG	ACGAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
EG327	AAGACAAAGG	CGAGAATGAT	TCTTCTACAG	ACAAAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA
Consensus	AAG-CAAAG-	CGAGAATG-T	TCTTCTACAG	AC-AAGGCGA	AGGCTTAGTG	ACTGCAAAAG	AAGTGATTGA

C5

	981						1050
H15	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
BZ10	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
BZ198	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
P20	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
H38	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
Z2491	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
H41	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
EG329	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
PMC21	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
EG327	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG
Consensus	TGCAGTAAAC	AAGGCTGGTT	GGAGAATGAA	AACAACAACC	GCTAATGGTC	AAACAGGTCA	AGCTGACAAG

C5

	1051						1120
H15	TTTGAAACCG	TTACATCAGG	CACAAAAGTA	ACCTTTGCTA	GTGGTAATGG	TACAACTGCG	ACTGTAAGTA
BZ10	TTTGAAACCG	TTACATCAGG	CACAAAAGTA	ACCTTTGCTA	GTGGTAATGG	TACAACTGCG	ACTGTAAGTA
BZ198	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
P20	TTTGAAACCG	TTACATCAGG	CACAAAAGTA	ACCTTTGCTA	GTGGTAATGG	TACAACTGCG	ACTGTAAGTA
H38	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
Z2491	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
H41	TTTGAAACCG	TTACATCAGG	CACAAAAGTA	ACCTTTGCTA	GTGGTAATGG	TACAACTGCG	ACTGTAAGTA
EG329	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
PMC21	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
EG327	TTTGAAACCG	TTACATCAGG	CACAAATGTA	ACCTTTGCTA	GTGGTAAAGG	TACAACTGCG	ACTGTAAGTA
Consensus	TTTGAAACCG	TTACATCAGG	CACAAA-GTA	ACCTTTGCTA	GTGGTAA-GG	TACAACTGCG	ACTGTAAGTA

C5

	1121						1190
H15	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
BZ10	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
BZ198	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
P20	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
H38	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
Z2491	AAGATGATCA	AGGCAACATC	ACTGTTATGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
H41	AAGATGATCA	AGGCAACATC	ACTGTTAAGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
EG329	AAGATGATCA	AGGCAACATC	ACTGTTATGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
PMC21	AAGATGATCA	AGGCAACATC	ACTGTTATGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
EG327	AAGATGATCA	AGGCAACATC	ACTGTTATGT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT
Consensus	AAGATGATCA	AGGCAACATC	ACTGTTA-GT	ATGATGTAAA	TGTCGGCGAT	GCCCTAAACG	TCAATCAGCT

C5

FIG. 2E

```
          1191                                                                  1260
      H15 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
     BZ10 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    BZ198 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      P20 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      H38 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    Z2491 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      H41 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    EG329 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    PMC21 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    EG327 GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
Consensus GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
                                          C5

          1261                                                                  1330
      H15 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
     BZ10 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    BZ198 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
      P20 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
      H38 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    Z2491 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTAGCC
      H41 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    EG329 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    PMC21 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    EG327 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
Consensus GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTA-CC
                                          C5

          1331                                                                  1400
      H15 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
     BZ10 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
    BZ198 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGGCGCC GCAGTTTTCC AGCGTTTCGC TCGGTGCGGG
      P20 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
      H38 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    Z2491 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGGCGCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
      H41 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
    EG329 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    PMC21 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    EG327 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
Consensus GCAACGG-AA AAATATCGAC ATCGCCACTT CGATG-C-CC GCA-TTTTCC AGCGTTTCGC TCGG-GCGGG
                                          C5

          1401                                                                  1470
      H15 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
     BZ10 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    BZ198 GGCGGATGCG CCCACTTTGA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TACCAACAAA
      P20 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
      H38 GGCGGATGCG CCCACTTTGA GCGTGGATGA CAAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    Z2491 GGCAGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
      H41 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    EG329 GGCGGATGCG CCCACTTTGA GCGTGGAT.. .GGGGACGCA TTGAATGTCG GCAGCAAGAA GGACAACAAA
    PMC21 GGCGGATGCG CCCACTTTGA GCGTGGAT.. .GGGGACGCA TTGAATGTCG GCAGCAAGAA GGACAACAAA
    EG327 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
Consensus GGC-GATGCG CCCACTTT-A GCGTGGAT-- ---GG-CGC- TTGAATGTCG GCAGCAAG-A ---CAACAAA
                                          C5
```

FIG. 2F

```
           1471                                                                1540
H15        CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
BZ10       CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
BZ198      CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
P20        CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
H38        CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
Z2491      CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
H41        CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCG CAACTTAAAG
EG329      CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
PMC21      CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
EG327      CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
Consensus  CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGC- CAACTTAAAG
                                         C5

           1541                                                                1610
H15        GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
BZ10       GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
BZ198      GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
P20        GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGAA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
H38        GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
Z2491      GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
H41        GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGAA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
EG329      GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
PMC21      GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
EG327      GCGTGGCGCA AAACTTGAAC AACCACATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
Consensus  G-GTGGCGCA AAACTTGAAC AACC-CATCG ACAATGTG-A CGGCAACGCG CG-GCGGG-A TCGCCCAAGC
                                         C5

           1611                                                                1680
H15        GATTGCAACC GCAGGTTTGG CTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
BZ10       GATTGCAACC GCAGGTTTGG CTCAGGCCTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
BZ198      GATTGCAACC GCAGGTCTAG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGACACT
P20        GATTGCAACC GCAGGTTTGG CTCAGGCCTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
H38        GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
Z2491      GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
H41        GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
EG329      GATTGCAACC GCAGGTCTGG TTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
PMC21      GATTGCAACC GCAGGTCTGG TTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
EG327      GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
Consensus  GATTGCAACC GCAGGT-T-G -TCAGGC-TA T-TGCCCGGC AAGAGTATGA TGGCGATCGG CGGCG--ACT
                                         C5
```

FIG. 2G

```
           1681                                                                         1750
H15        TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
BZ10       TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
BZ198      TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCAAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
P20        TATCTCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
H38        TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
Z2491      TATCTCGGCG AAGCCGGTTA TGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
H41        TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCAAGCA TTTCCGCCGG CGGAAATTGG ATTATCAAAG
EG329      TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
PMC21      TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
EG327      TATCGCGGCG AAGCCGGTTA TGCCATCGGC TACTCAAGCA TTTCCGACGG CGGAAATTGG ATTATCAAAG
Consensus  TATC-CGGCG AAGCCGGTTA -GCCATCGGC TACTC-AG-A TTTC-G-C-- -GG-AATTGG -TTATCAA-G
                                             C5

           1751                                                                1815
H15        GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
BZ10       GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTACTTC CGCATCTGTC GGTTATCAGT GGTAA
BZ198      GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
P20        GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTGCTTC CGCATCTGTC GGTTATCAAT GGTAA
H38        GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTACTTC CGCATCTGTC GGTTATCAGT GGTAA
Z2491      GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
H41        GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
EG329      GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
PMC21      GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
EG327      GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
Consensus  GCACGGCTTC CGGCAATTCG CGCGG-CATT TCGGT-CTTC CGCATCTGTC GGTTATCA-T GGTAA
                                             C5
```

FIG. 2H

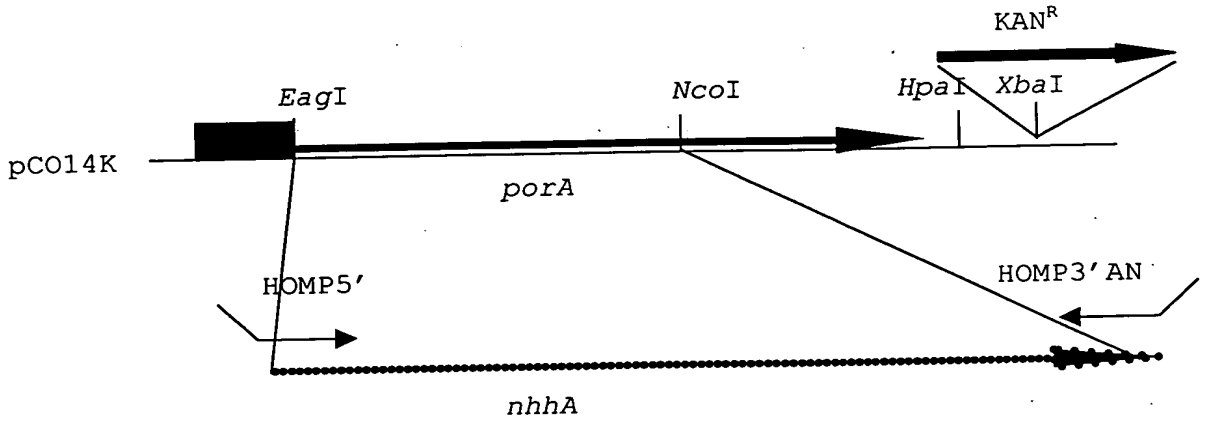

FIG. 3A

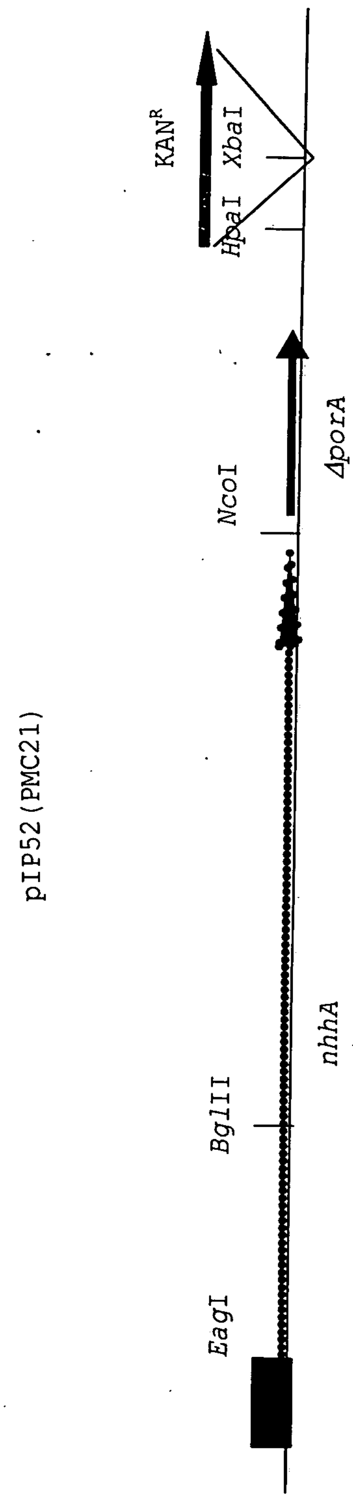

FIG. 3B

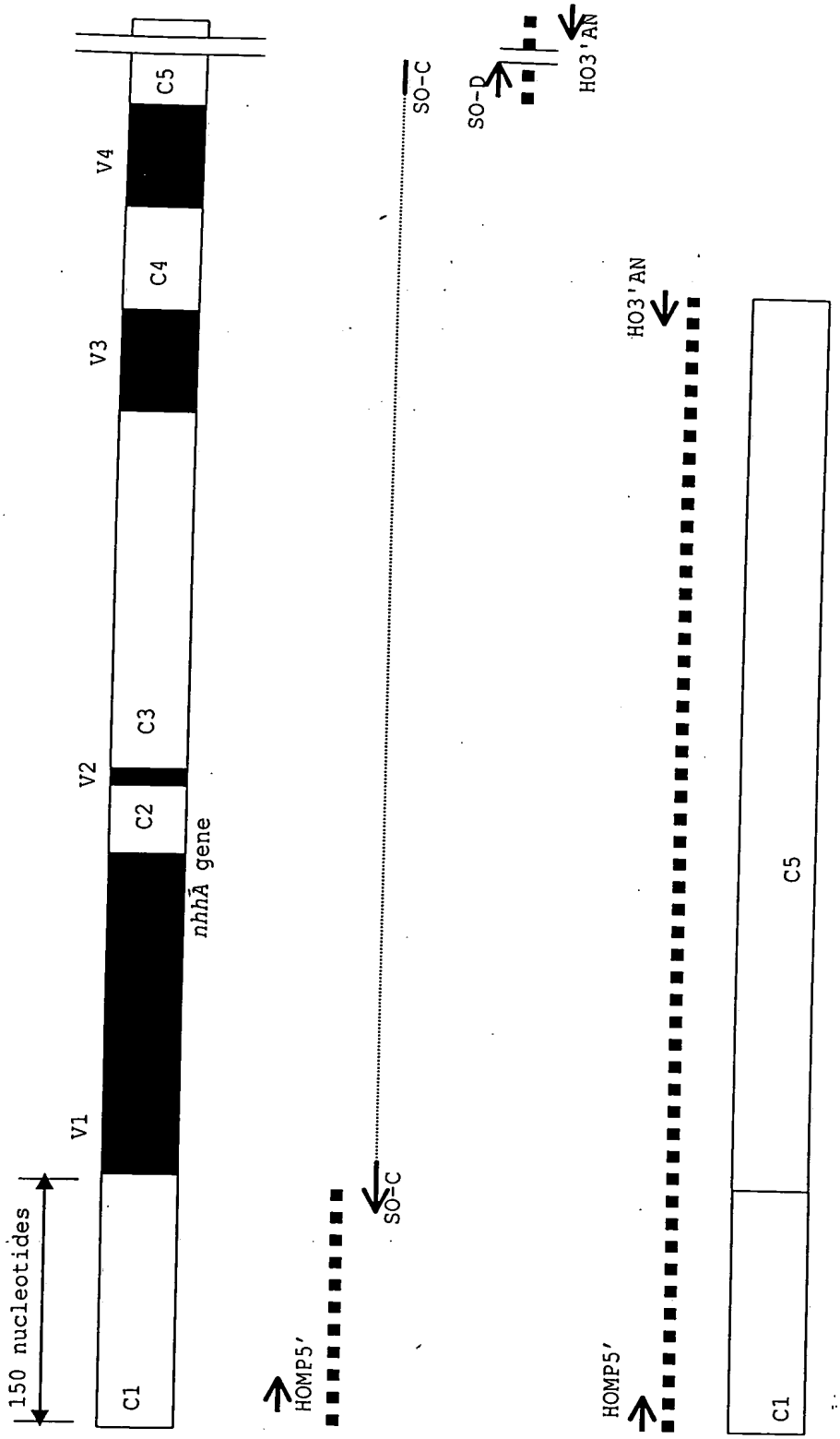

FIG. 4A

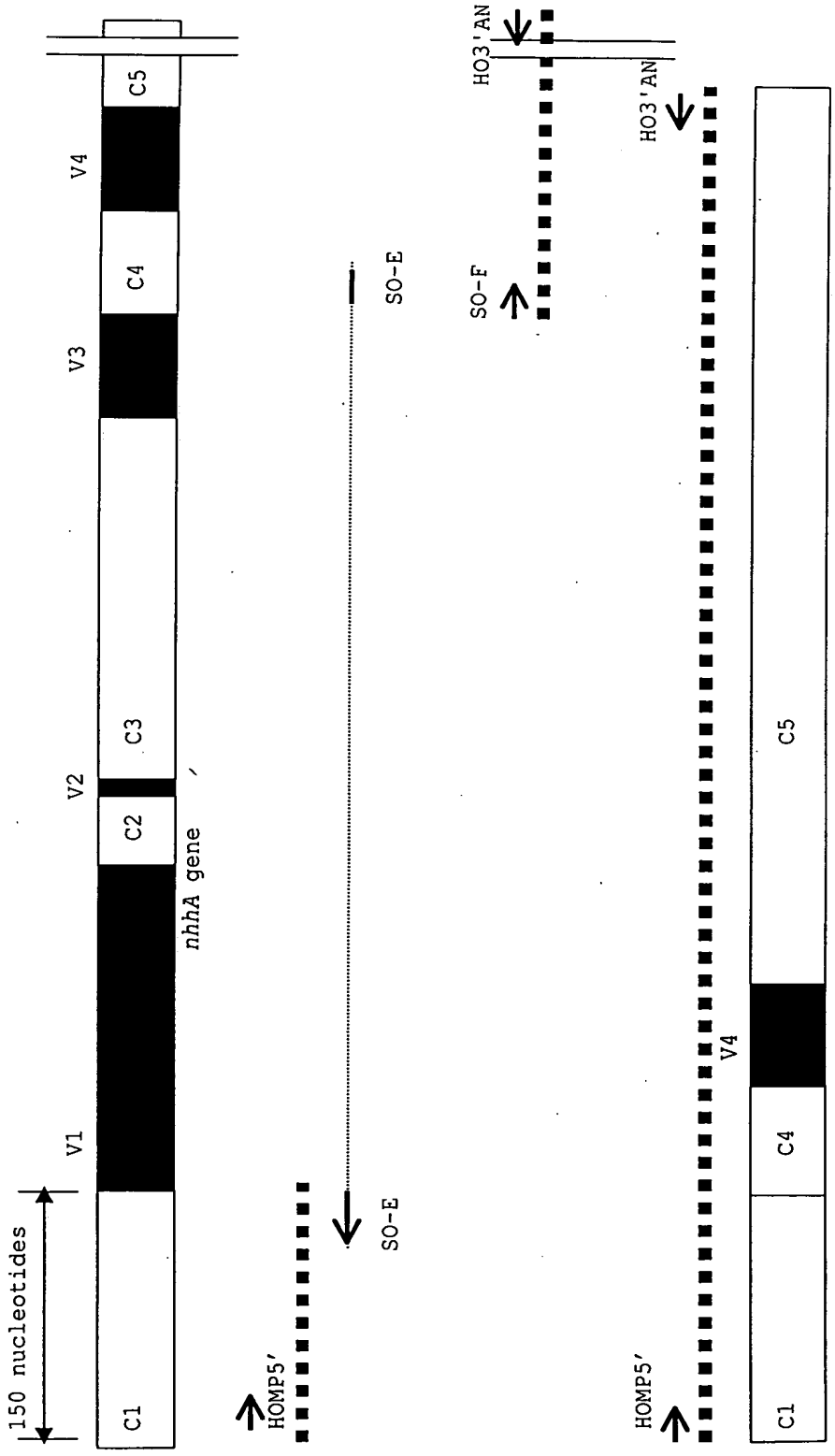
C5
V4
C4
V3
C3
V2
C2
nhhA gene
V1
C1
150 nucleotides
HOMP5'
SO-E
SO-E
SO-F
HO3'AN
HO3'AN
HOMP5'
V4
C4
C5
C1

FIG. 4B

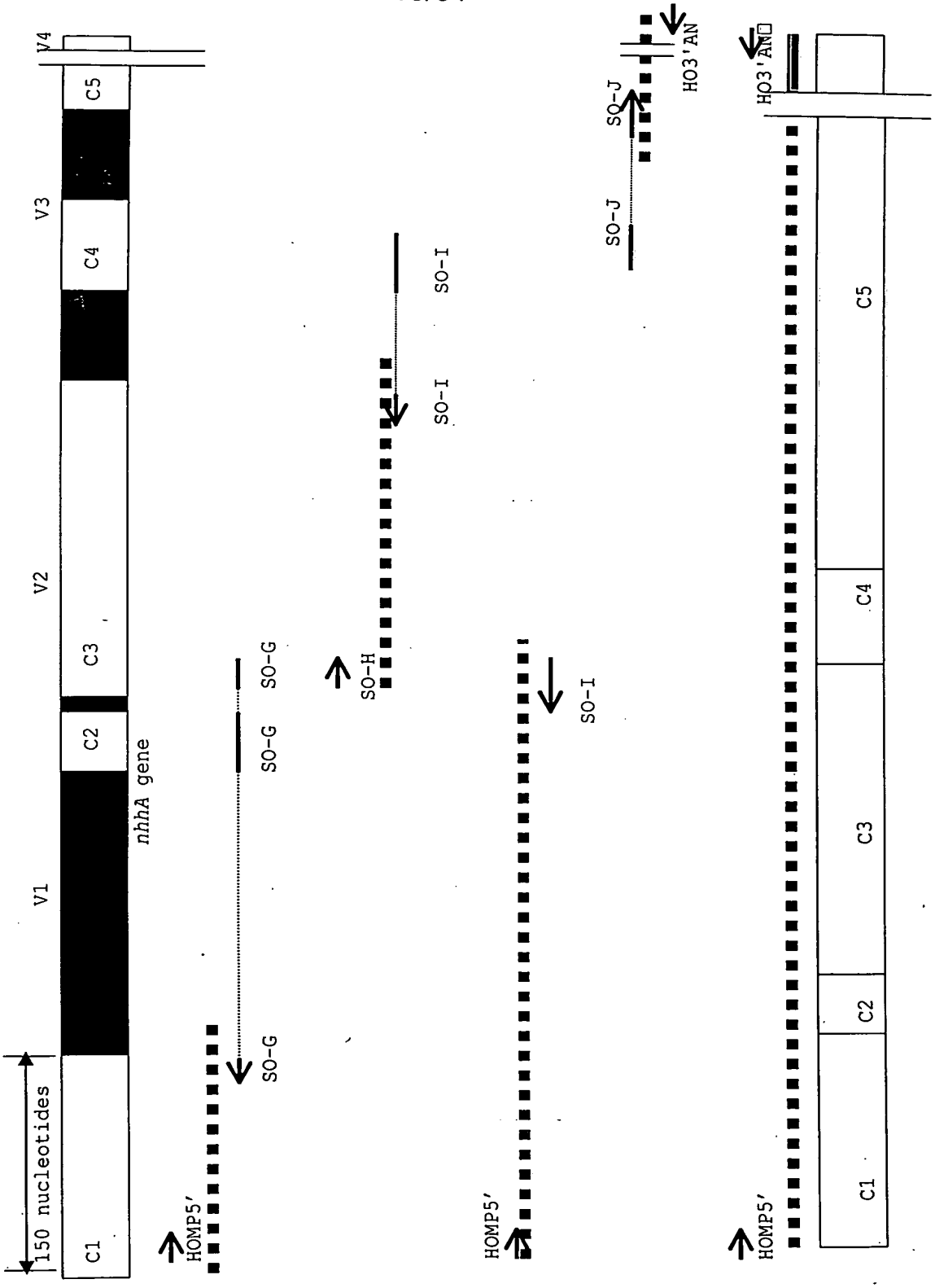
V4
C5
V3
C4
V2
C3
C2
V1
C1
nhhA gene
150 nucleotides
HOMP5'
SO-G
SO-G
SO-G
SO-H
SO-I
SO-I
SO-I
HOMP5'
SO-J
SO-J
HO3'AN
HO3'AN
C5
C4
C3
C2
C1
HOMP5'

FIG. 4C

```
  1  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51  ANNETDLTSV GTEKLSFSAN GNKVNITSDT KGLNFAKETA GTNGDTTVHL
101  NGIGSTLTDT LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151  PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201  VIKEKDGKLV TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251  NGQTGQADKF ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA
301  LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351  ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV
401  RITNVAPGVK EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG
451  LVQAYLPGKS MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG
501  HFGASASVGY QW*
```

FIG. 5A

```
   1  ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101  TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151  GCTAACAATG AAACAGATCT GACCAGTGTT GGAACTGAAA AATTATCGTT
 201  TAGCGCAAAC GGCAATAAAG TCAACATCAC AAGCGACACC AAAGGCTTGA
 251  ATTTTGCGAA AGAAACGGCT GGGACGAACG GCGACACCAC GGTTCATCTG
 301  AACGGTATTG GTTCGACTTT GACCGATACG CTGCTGAATA CCGGAGCGAC
 351  CACAAACGTA ACCAACGACA ACGTTACCGA TGACGAGAAA AAACGTGCGG
 401  CAAGCGTTAA AGACGTATTA AACGCTGGCT GGAACATTAA AGGCGTTAAA
 451  CCCGGTACAA CAGCTTCCGA TAACGTTGAT TTCGTCCGCA CTTACGACAC
 501  AGTCGAGTTC TTGAGCGCAG ATACGAAAAC AACGACTGTT AATGTGGAAA
 551  GCAAAGACAA CGGCAAGAAA ACCGAAGTTA AAATCGGTGC GAAGACTTCT
 601  GTTATTAAAG AAAAAGACGG TAAGTTGGTT ACTGGTAAAG ACAAAGGCGA
 651  GAATGGTTCT TCTACAGACG AAGGCGAAGG CTTAGTGACT GCAAAAGAAG
 701  TGATTGATGC AGTAAACAAG GCTGGTTGGA GAATGAAAAC AACAACCGCT
 751  AATGGTCAAA CAGGTCAAGC TGACAAGTTT GAAACCGTTA CATCAGGCAC
 801  AAATGTAACC TTTGCTAGTG GTAAAGGTAC AACTGCGACT GTAAGTAAAG
 851  ATGATCAAGG CAACATCACT GTTATGTATG ATGTAAATGT CGGCGATGCC
 901  CTAAACGTCA ATCAGCTGCA AAACAGCGGT TGGAATTTGG ATTCCAAAGC
 951  GGTTGCAGGT TCTTCGGGCA AAGTCATCAG CGGCAATGTT TCGCCGAGCA
1001  AGGGAAAGAT GGATGAAACC GTCAACATTA ATGCCGGCAA CAACATCGAG
1051  ATTACCCGCA ACGGTAAAAA TATCGACATC GCCACTTCGA TGACCCCGCA
1101  GTTTTCCAGC GTTTCGCTCG GCGCGGGGGC GGATGCGCCC ACTTTGAGCG
1151  TGGATGGGGA CGCATTGAAT GTCGGCAGCA AGAAGGACAA CAAACCCGTC
1201  CGCATTACCA ATGTCGCCCC GGGCGTTAAA GAGGGGGATG TTACAAACGT
1251  CGCACAACTT AAAGGCGTGG CGCAAAACTT GAACAACCGC ATCGACAATG
1301  TGGACGGCAA CGCGCGTGCG GGCATCGCCC AAGCGATTGC AACCGCAGGT
1351  CTGGTTCAGG CGTATTTGCC CGGCAAGAGT ATGATGGCGA TCGGCGGCGG
1401  CACTTATCGC GGCGAAGCCG GTTACGCCAT CGGCTACTCC AGTATTTCCG
1451  ACGGCGGAAA TTGGATTATC AAAGGCACGG CTTCCGGCAA TTCGCGCGGC
1501  CATTTCGGTG CTTCCGCATC TGTCGGTTAT CAGTGGTAA
```

FIG. 5B

1 MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
51 ATDETGLINV ETEKLSFGAN GKKVNIISDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDM LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKGKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTKVT FASGNGTTAT VSKDDQGNIT VKYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDDEGAL NVGSKDANKP
401 VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN RIDNVNGNAR AGIAQAIATA
451 GLVQAYLPGK SMMAIGGGTY LGEAGYAIGY SSISAGGNWI IKGTASGNSR
501 GHFGASASVG YQW*

FIG. 6A

1 ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT
51 CGCCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
101 TGAAGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT TCAGGCGAAT
151 GCTACCGATG AAACAGGCCT GATCAATGTT GAAACTGAAA AATTATCGTT
201 TGGCGCAAAC GGCAAGAAAG TCAACATCAT AAGCGACACC AAAGGCTTGA
251 ATTTCGCGAA AGAAACGGCT GGGACGAACG GCGACACCAC GGTTCATCTG
301 AACGGTATCG GTTCGACTTT GACCGATATG CTGCTGAATA CCGGAGCGAC
351 CACAAACGTA ACCAACGACA ACGTTACCGA TGACGAGAAA AAACGTGCGG
401 CAAGCGTTAA AGACGTATTA AACGCAGGCT GGAACATTAA AGGCGTTAAA
451 CCCGGTACAA CAGCTTCCGA TAACGTTGAT TTCGTCCGCA CTTACGACAC
501 AGTCGAGTTC TTGAGCGCAG ATACGAAAAC AACGACTGTT AATGTGGAAA
551 GCAAAGACAA CGGCAAGAAA ACCGAAGTTA AAATCGGTGC GAAGACTTCT
601 GTTATTAAAG AAAAAGACGG TAAGTTGGTT ACTGGTAAAG GCAAAGGCGA
651 GAATGGTTCT TCTACAGACG AAGGCGAAGG CTTAGTGACT GCAAAAGAAG
701 TGATTGATGC AGTAAACAAG GCTGGTTGGA GAATGAAAAC AACAACCGCT
751 AATGGTCAAA CAGGTCAAGC TGACAAGTTT GAAACCGTTA CATCAGGCAC
801 AAAAGTAACC TTTGCTAGTG GTAATGGTAC AACTGCGACT GTAAGTAAAG
851 ATGATCAAGG CAACATCACT GTTAAGTATG ATGTAAATGT CGGCGATGCC
901 CTAAACGTCA ATCAGCTGCA AAACAGCGGT TGGAATTTGG ATTCCAAAGC
951 GGTTGCAGGT TCTTCGGGCA AAGTCATCAG CGGCAATGTT TCGCCGAGCA
1001 AGGGAAAGAT GGATGAAACC GTCAACATTA ATGCCGGCAA CAACATCGAG
1051 ATTACCCGCA ACGGCAAAAA TATCGACATC GCCACTTCGA TGACCCCGCA
1101 ATTTTCCAGC GTTTCGCTCG GCGCGGGGGC GGATGCGCCC ACTTTAAGCG
1151 TGGATGACGA GGGCGCGTTG AATGTCGGCA GCAAGGATGC CAACAAACCC
1201 GTCCGCATTA CCAATGTCGC CCCGGGCGTT AAAGAGGGGG ATGTTACAAA
1251 CGTCGCGCAA CTTAAAGGTG TGGCGCAAAA CTTGAACAAC CGCATCGACA
1301 ATGTGAACGG CAACGCGCGT GCGGGCATCG CCCAAGCGAT TGCAACCGCA
1351 GGTCTGGTTC AGGCGTATCT GCCCGGCAAG AGTATGATGG CGATCGGCGG
1401 CGGCACTTAT CTCGGCGAAG CCGGTTATGC CATCGGCTAC TCAAGCATTT
1451 CCGCCGGCGG AAATTGGATT ATCAAAGGCA CGGCTTCCGG CAATTCGCGC
1501 GGCCATTTCG GTGCTTCCGC ATCTGTCGGT TATCAGTGGT AA

FIG. 6B

```
  1  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51  ANNVDFVRTY DTVEFLSADT KTTTVNVESK DNGKKTEVKI GAKTSVIKEK
101  DGKLVTGKDK GENGSSTDEG EGLVTAKEVI DAVNKAGWRM KTTTANGQTG
151  QADKFETVTS GTNVTFASGK GTTATVSKDD QGNITVMYDV NVGDALNVNQ
201  LQNSGWNLDS KAVAGSSGKV ISGNVSPSKG KMDETVNINA GNNIEITRNG
251  KNIDIATSMT PQFSSVSLGA GADAPTLSVD GDALNVGSKK DNKPVRITNV
301  APGVKEGDVT NVAQLKGVAQ NLNNRIDNVD GNARAGIAQA IATAGLVQAY
351  LPGKSMMAIG GGTYRGEAGY AIGYSSISDG GNWIIKGTAS GNSRGHFGAS
401  ASVGYQW*
```

FIG. 7A

```
   1  ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101  TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151  GCTAACAACG TTGATTTCGT CCGCACTTAC GACACAGTCG AGTTCTTGAG
 201  CGCAGATACG AAAACAACGA CTGTTAATGT GGAAAGCAAA GACAACGGCA
 251  AGAAAACCGA AGTTAAAATC GGTGCGAAGA CTTCTGTTAT TAAAGAAAAA
 301  GACGGTAAGT TGGTTACTGG TAAAGACAAA GGCGAGAATG GTTCTTCTAC
 351  AGACGAAGGC GAAGGCTTAG TGACTGCAAA AGAAGTGATT GATGCAGTAA
 401  ACAAGGCTGG TTGGAGAATG AAAACAACAA CCGCTAATGG TCAAACAGGT
 451  CAAGCTGACA AGTTTGAAAC CGTTACATCA GGCACAAATG TAACCTTTGC
 501  TAGTGGTAAA GGTACAACTG CGACTGTAAG TAAAGATGAT CAAGGCAACA
 551  TCACTGTTAT GTATGATGTA AATGTCGGCG ATGCCCTAAA CGTCAATCAG
 601  CTGCAAAACA GCGGTTGGAA TTTGGATTCC AAAGCGGTTG CAGGTTCTTC
 651  GGGCAAAGTC ATCAGCGGCA ATGTTTCGCC GAGCAAGGGA AAGATGGATG
 701  AAACCGTCAA CATTAATGCC GGCAACAACA TCGAGATTAC CCGCAACGGT
 751  AAAAATATCG ACATCGCCAC TTCGATGACC CCGCAGTTTT CCAGCGTTTC
 801  GCTCGGCGCG GGGGCGGATG CGCCCACTTT GAGCGTGGAT GGGGACGCAT
 851  TGAATGTCGG CAGCAAGAAG GACAACAAAC CCGTCCGCAT TACCAATGTC
 901  GCCCCGGGCG TTAAAGAGGG GGATGTTACA AACGTCGCAC AACTTAAAGG
 951  CGTGGCGCAA AACTTGAACA ACCGCATCGA CAATGTGGAC GGCAACGCGC
1001  GTGCGGGCAT CGCCCAAGCG ATTGCAACCG CAGGTCTGGT TCAGGCGTAT
1051  TTGCCCGGCA AGAGTATGAT GGCGATCGGC GGCGGCACTT ATCGCGGCGA
1101  AGCCGGTTAC GCCATCGGCT ACTCCAGTAT TTCCGACGGC GGAAATTGGA
1151  TTATCAAAGG CACGGCTTCC GGCAATTCGC GCGGCCATTT CGGTGCTTCC
1201  GCATCTGTCG GTTATCAGTG GTAA
```

FIG. 7B

```
  1 MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51 ANRAASVKDV LNAGWNIKGV KPGTTASDNV DFVRTYDTVE FLSADTKTTT
101 VNVESKDNGK KTEVKIGAKT SVIKEKDGKL VTGKDKGENG SSTDEGEGLV
151 TAKEVIDAVN KAGWRMKTTT ANGQTGQADK FETVTSGTNV TFASGKGTTA
201 TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS GWNLDSKAVA GSSGKVISGN
251 VSPSKGKMDE TVNINAGNNI EITRNGKNID IATSMTPQFS SVSLGAGADA
301 PTLSVDGDAL NVGSKKDNKP VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN
351 RIDNVDGNAR AGIAQAIATA GLVQAYLPGK SMMAIGGGTY RGEAGYAIGY
401 SSISDGGNWI IKGTASGNSR GHFGASASVG YQW*
```

FIG. 8A

```
   1 ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51 CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101 TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151 GCTAACCGTG CGGCAAGCGT TAAAGACGTA TTAAACGCTG GCTGGAACAT
 201 TAAAGGCGTT AAACCCGGTA CAACAGCTTC CGATAACGTT GATTTCGTCC
 251 GCACTTACGA CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT
 301 GTTAATGTGG AAAGCAAAGA CAACGGCAAG AAAACCGAAG TTAAAATCGG
 351 TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
 401 AAGACAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG
 451 ACTGCAAAAG AAGTGATTGA TGCAGTAAAC AAGGCTGGTT GGAGAATGAA
 501 AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG TTTGAAACCG
 551 TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG
 601 ACTGTAAGTA AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA
 651 TGTCGGCGAT GCCCTAAACG TCAATCAGCT GCAAAACAGC GGTTGGAATT
 701 TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
 751 GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG
 801 CAACAACATC GAGATTACCC GCAACGGTAA AAATATCGAC ATCGCCACTT
 851 CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG GGCGGATGCG
 901 CCCACTTTGA GCGTGGATGG GGACGCATTG AATGTCGGCA GCAAGAAGGA
 951 CAACAAACCC GTCCGCATTA CCAATGTCGC CCCGGGCGTT AAAGAGGGGG
1001 ATGTTACAAA CGTCGCACAA CTTAAAGGCG TGGCGCAAAA CTTGAACAAC
1051 CGCATCGACA ATGTGGACGG CAACGCGCGT GCGGGCATCG CCCAAGCGAT
1101 TGCAACCGCA GGTCTGGTTC AGGCGTATTT GCCCGGCAAG AGTATGATGG
1151 CGATCGGCGG CGGCACTTAT CGCGGCGAAG CCGGTTACGC CATCGGCTAC
1201 TCCAGTATTT CCGACGGCGG AAATTGGATT ATCAAAGGCA CGGCTTCCGG
1251 CAATTCGCGC GGCCATTTCG GTGCTTCCGC ATCTGTCGGT TATCAGTGGT
1301 AA
```

FIG. 8B

1 MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
51 ANTLKAGDNL KIKQFTYSLK KDLTDLTSVG TEKLSFSANG NKVNITSDTK
101 GLNFAKETAG TNGDTTVHLN GIGSTLTDRA ASVKDVLNAG WNIKGVKNVD
151 FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS VIKEKDGKLV
201 TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA NGQTGQADKF
251 ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA LNVNQLQNSG
301 WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE ITRNGKNIDI
351 ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV RITNVAPGVK
401 EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG LVQAYLPGKS
451 MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG HFGASASVGY
501 QW*

FIG. 9A

1 ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
51 CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
101 TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
151 GCTAACACCC TCAAAGCCGG CGACAACCTG AAAATCAAAC AATTCACCTA
201 CTCGCTGAAA AAAGACCTCA CAGATCTGAC CAGTGTTGGA ACTGAAAAAT
251 TATCGTTTAG CGCAAACGGC AATAAAGTCA ACATCACAAG CGACACCAAA
301 GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
351 TCATCTGAAC GGTATTGGTT CGACTTTGAC CGATCGTGCG GCAAGCGTTA
401 AAGACGTATT AAACGCTGGC TGGAACATTA AAGGCGTTAA AAACGTTGAT
451 TTCGTCCGCA CTTACGACAC AGTCGAGTTC TTGAGCGCAG ATACGAAAAC
501 AACGACTGTT AATGTGGAAA GCAAAGACAA CGGCAAGAAA ACCGAAGTTA
551 AAATCGGTGC GAAGACTTCT GTTATTAAAG AAAAAGACGG TAAGTTGGTT
601 ACTGGTAAAG ACAAAGGCGA GAATGGTTCT TCTACAGACG AAGGCGAAGG
651 CTTAGTGACT GCAAAAGAAG TGATTGATGC AGTAAACAAG GCTGGTTGGA
701 GAATGAAAAC AACAACCGCT AATGGTCAAA CAGGTCAAGC TGACAAGTTT
751 GAAACCGTTA CATCAGGCAC AAATGTAACC TTTGCTAGTG GTAAAGGTAC
801 AACTGCGACT GTAAGTAAAG ATGATCAAGG CAACATCACT GTTATGTATG
851 ATGTAAATGT CGGCGATGCC CTAAACGTCA ATCAGCTGCA AAACAGCGGT
901 TGGAATTTGG ATTCCAAAGC GGTTGCAGGT TCTTCGGGCA AAGTCATCAG
951 CGGCAATGTT TCGCCGAGCA AGGGAAAGAT GGATGAAACC GTCAACATTA
1001 ATGCCGGCAA CAACATCGAG ATTACCCGCA ACGGTAAAAA TATCGACATC
1051 GCCACTTCGA TGACCCCGCA GTTTTCCAGC GTTTCGCTCG GCGCGGGGGC
1101 GGATGCGCCC ACTTTGAGCG TGGATGGGGA CGCATTGAAT GTCGGCAGCA
1151 AGAAGGACAA CAAACCCGTC CGCATTACCA ATGTCGCCCC GGGCGTTAAA
1201 GAGGGGGATG TTACAAACGT CGCACAACTT AAAGGCGTGG CGCAAAACTT
1251 GAACAACCGC ATCGACAATG TGGACGGCAA CGCGCGTGCG GGCATCGCCC
1301 AAGCGATTGC AACCGCAGGT CTGGTTCAGG CGTATTTGCC CGGCAAGAGT
1351 ATGATGGCGA TCGGCGGCGG CACTTATCGC GGCGAAGCCG GTTACGCCAT
1401 CGGCTACTCC AGTATTTCCG ACGGCGGAAA TTGGATTATC AAAGGCACGG
1451 CTTCCGGCAA TTCGCGCGGC CATTTCGGTG CTTCCGCATC TGTCGGTTAT
1501 CAGTGGTAA

FIG. 9B

```
                1                                                   50
H41             MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
PMC21           MNKIYRIIWN SALNAWVVVS DLTRNHTKPA SATVNTAVLA TLLFATVQAS
H41Studel       MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
PMC21Bgldel     MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
PMC21C1C5       MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
                                      C1

                51                                                  100
H41             ATDED..EEE ELESVQRS.V VGSIQASMEG SVELET...I SLSMTNDSKE
PMC21           ANNEEQEEYL YLHPVQRTVA VLIVNSDKEG AGEKEKVEEN SDWAVYFNEK
H41Studel       ATDE...... .......... .......... .......... ..........
PMC21Bgldel     ANNE...... .......... .......... .......... ..........
PMC21C1C5       AN ....... .......... .......... .......... ..........
                V1

                101                                                 150
H41             FVDPYIVVTL KAGDNLKIKQ N.TNENTNAS SFTYSLKKDL TGLINVETEK
PMC21           GVLTAREITL KAGDNLKIKQ NGTN...... .FTYSLKKDL TDLTSVGTEK
H41Studel       .......... .......... .......... .......... TGLINVETEK
PMC21Bgldel     .......... .......... .......... .......... TDLTSVGTEK
PMC21C1C5       .......... .......... .......... .......... ..........
                V1         C2         V2                    C3

                151                                                 200
H41             LSFGANGKKV NIISDTKGLN FAKETAGTNG DTTVHLNGIG STLTDMLLNT
PMC21           LSFSAHGNKV NITSDTKGLN FAKETAGTNG DTTVHLNGIG STLTDTLLNT
H41Studel       LSFGANGKKV NIISDTKGLN FAKETAGTNG DTTVHLNGIG STLTDMLLNT
PMC21Bgldel     LSFSANGNKV NITSDTKGLN FAKETAGTNG DTTVHLNGIG STLTDTLLNT
PMC21C1C5       .......... .......... .......... .......... ..........
                                C3                                  V3

                201                                                 250
H41             GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21           GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
H41Studel       GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21Bgldel     GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21C1C5       .......... .......... .......... .......... ...NVDFVRT
                      V3                C4                  V4    C5

                251                                                 300
H41             YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKG
PMC21           YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
H41Studel       YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKG
PMC21Bgldel     YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
PMC21C1C5       YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
                                      C5

                301                                                 350
H41             KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21           KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
H41Studel       KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21Bgldel     KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21C1C5       KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
                                      C5

                351                                                 400
H41             SGTKVTFASE NGTTATVSKD DQGNITVKYD VNVGDALNVN QLQNSGWNLD
PMC21           SGTNVTFASE KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
H41Studel       SGTKVTFASE NGTTATVSKD DQGNITVKYD VNVGDALNVN QLQNSGWNLD
PMC21Bgldel     SGTNVTFASE KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
PMC21C1C5       SGTNVTFASE KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
                                      C5
```

FIG. 10A

```
            401                                                450
H41         SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
PMC21       SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
H41Studel   SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
PMC21Bgldel SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
PMC21C1C5   SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
                                   C5

            451                                                500
H41         TPQFSSVSLG AGADAPTLSV DDEGALNVGS KDANKPVRIT NVAPGVKEGD
PMC21       TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
H41Studel   TPQFSSVSLG AGADAPTLSV DDEGALNVGS KDANKPVRIT NVAPGVKEGD
PMC21Bgldel TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
PMC21C1C5   TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
                                   C5

            501                                                550
H41         VTNVAQLKGV AQNLNNRIDN VNGNARAGIA QAIATAGLVQ AYLPGKSMMA
PMC21       VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
H41Studel   VTNVAQLKGV AQNLNNRIDN VNGNARAGIA QAIATAGLVQ AYLPGKSMMA
PMC21Bgldel VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
PMC21C1C5   VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
                                   C5

            551                                                600
H41         IGGGTYLGEA GYAIGYSSIS AGGNWIIKGT ASGNSRGHFG ASASVGYQW.
PMC21       IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
H41Studel   IGGGTYLGEA GYAIGYSSIS AGGNWIIKGT ASGNSRGHFG ASASVGYQW.
PMC21Bgldel IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
PMC21C1C5   IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
                                   C5
```

FIG. 10B

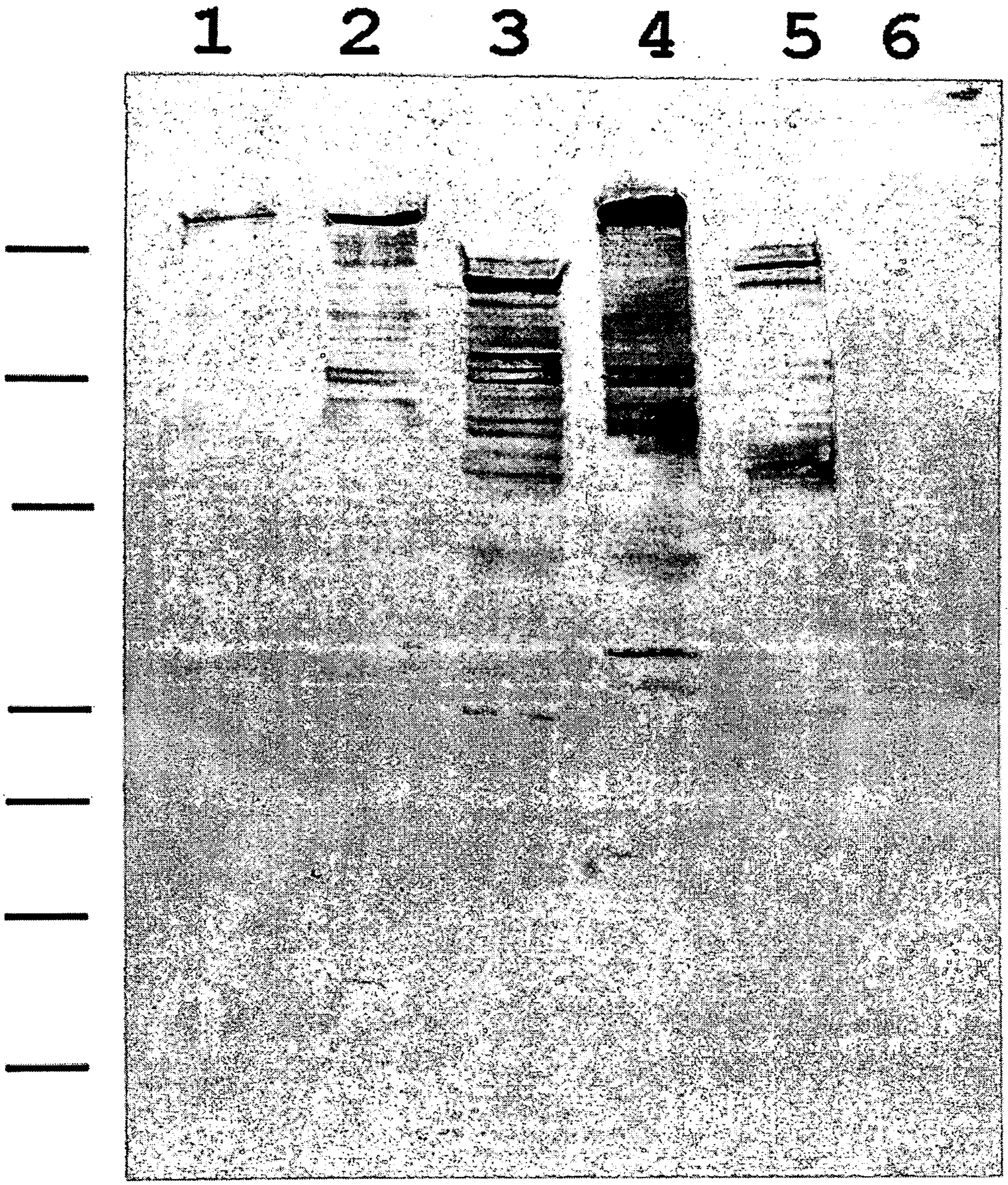

FIG. 11

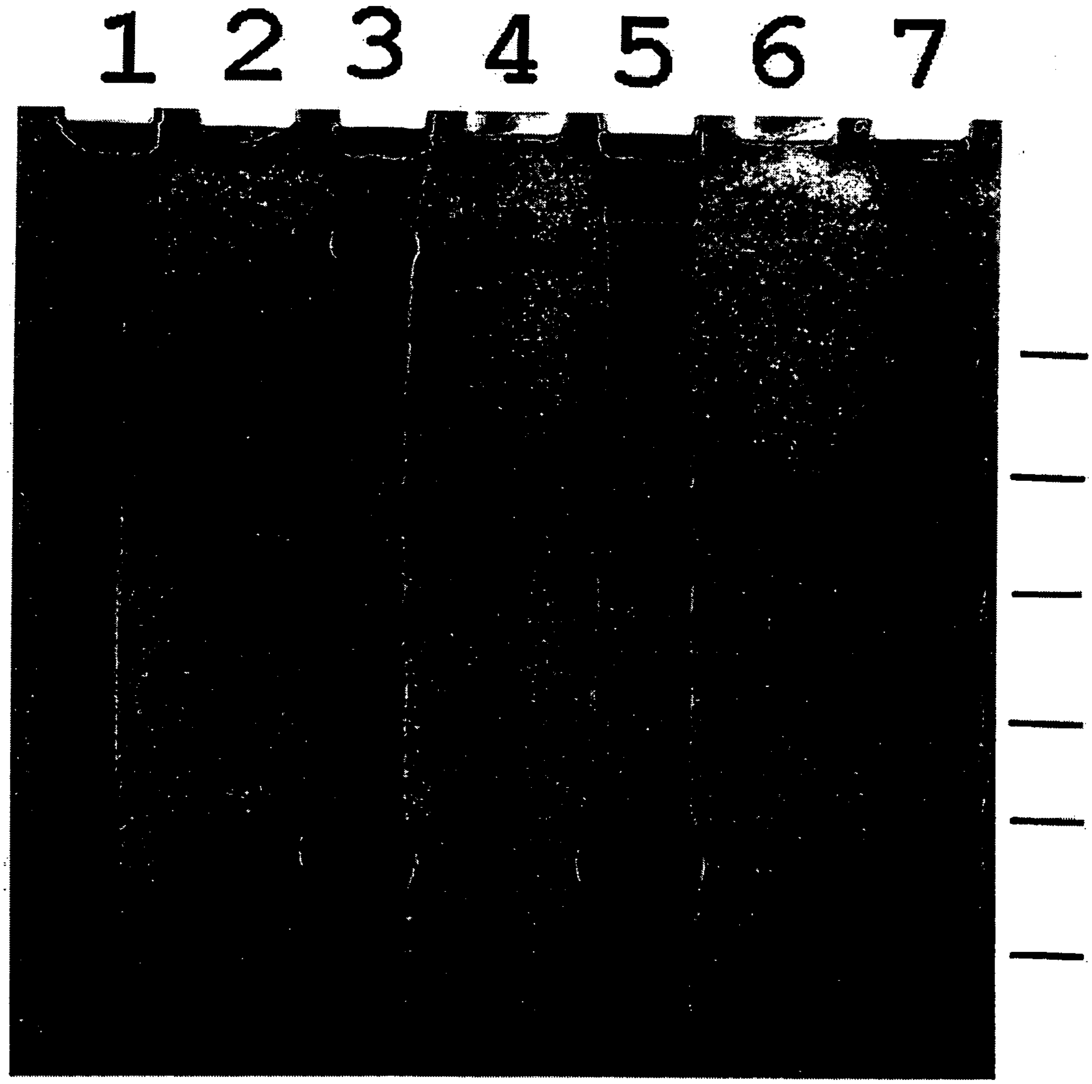

FIG. 12

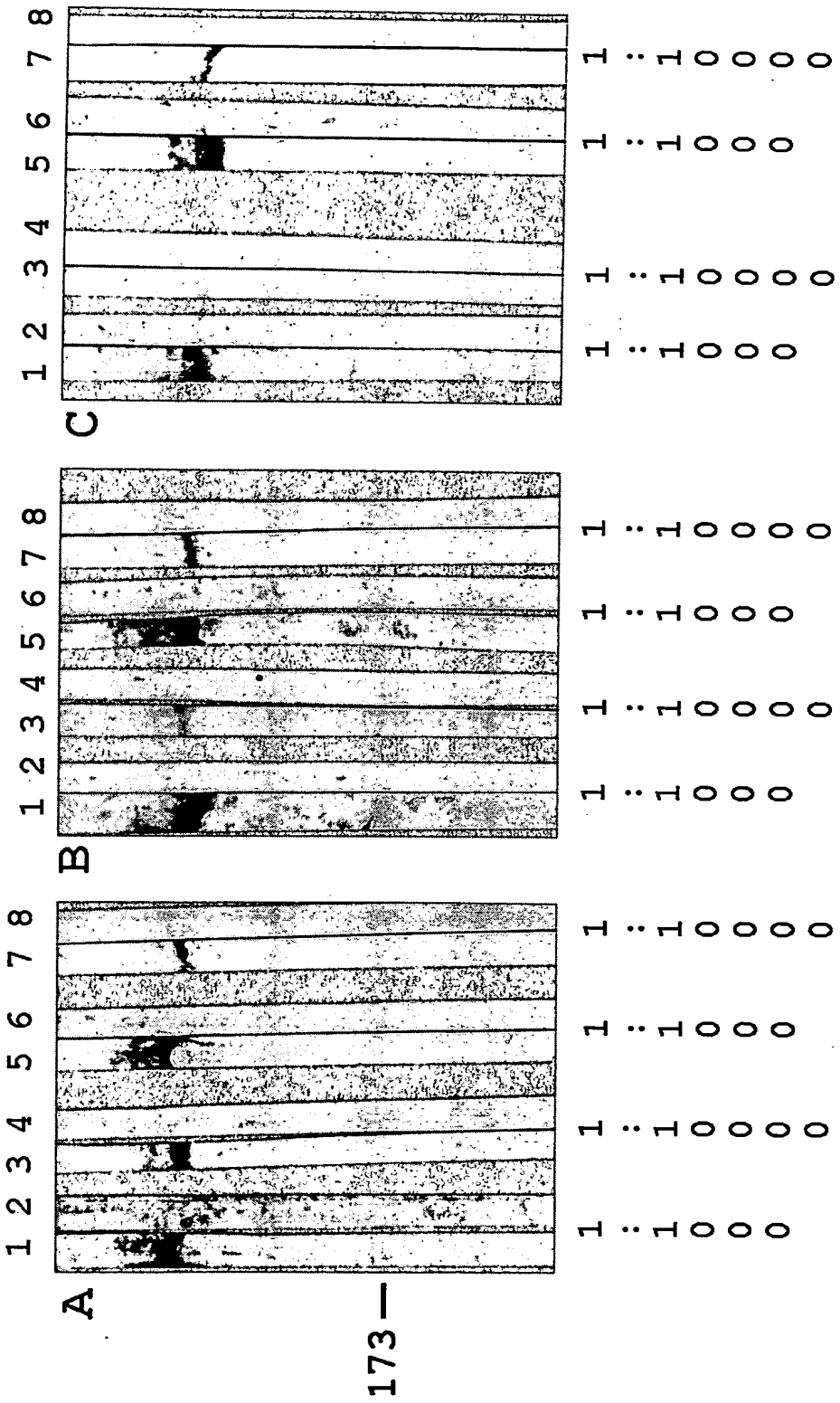
1 2 3 4 5 6 7 8
C
1 :: 10000
1 :: 1000
1 :: 10000
1 :: 1000
1 2 3 4 5 6 7 8
B
1 :: 10000
1 :: 1000
1 :: 10000
1 :: 1000
1 2 3 4 5 6 7 8
A
1 :: 10000
1 :: 1000
1 :: 10000
1 :: 1000
173

FIG. 13

```
52   NNEEQEEYL YLHPVQRTVA VLIVNSDKEG AGEKEKVEEN SDWAVYFNEK
101  GVLTAREITL KAGDNLKIKQ NGTNFTYSLK KDLTDLTSVG TEKLSFSAHG
151  NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL LNTGATTNVT
201  NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF VRTYDTVEFL
251  SADTKTTTVN VESKDNGKKT EVKIGAKTSV IKEKDGKLVT GKDKGENGSS
301  TDEGEGLVTA KEVIDAVNKA GWRMKTTTAN GQTGQADKFE TVTSGTNVTF
351  ASGKGTTATV SKDDQGNITV MYDVNVGDAL NVNQLQNSGW NLDSKAVAGS
401  SGKVISGNVS PSKGKMDETV NINAGNNIEI TRNGKNIDIA TSMTPQFSSV
451  SLGAGADAPT LSVDGDALNV GSKKDNKPVR ITNVAPGVKE GDVTNVAQLK
501  GVAQNLNNRI DNVDGNARAG IAQAIATAGL VQAYLPGKSM MAIGGGTYRG
551  EAGYAIGYSS ISDGGNWIIK GTASGNSRGH FGASASVGYQ W*
```

FIG. 14A

```
52   TDEDEEEEL ESVQRSVVGS IQASMEGSVE LETISLSMTN DSKEFVDPYI
101  VVTLKAGDNL KIKQNTNENT NASSFTYSLK KDLTGLINVE TEKLSFGANG
151  KKVNIISDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDML LNTGATTNVT
201  NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF VRTYDTVEFL
251  SADTKTTTVN VESKDNGKKT EVKIGAKTSV IKEKDGKLVT GKGKGENGSS
301  TDEGEGLVTA KEVIDAVNKA GWRMKTTTAN GQTGQADKFE TVTSGTKVTF
351  ASGNGTTATV SKDDQGNITV KYDVNVGDAL NVNQLQNSGW NLDSKAVAGS
401  SGKVISGNVS PSKGKMDETV NINAGNNIEI TRNGKNIDIA TSMTPQFSSV
451  SLGAGADAPT LSVDDEGALN VGSKDANKPV RITNVAPGVK EGDVTNVAQL
501  KGVAQNLNNR IDNVNGNARA GIAQAIATAG LVQAYLPGKS MMAIGGGTYL
551  GEAGYAIGYS SISAGGNWII KGTASGNSRG HFGASASVGY QW*
```

FIG. 14B

52 NNETDLTSV GTEKLSFSAN GNKVNITSDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDT LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV
401 RITNVAPGVK EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG
451 LVQAYLPGKS MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG
501 HFGASASVGY QW*

FIG. 14C

52 TDETGLINV ETEKLSFGAN GKKVNIISDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDM LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKGKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTKVT FASGNGTTAT VSKDDQGNIT VKYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDDEGAL NVGSKDANKP
401 VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN RIDNVNGNAR AGIAQAIATA
451 GLVQAYLPGK SMMAIGGGTY LGEAGYAIGY SSISAGGNWI IKGTASGNSR
501 GHFGASASVG YQW*

FIG. 14D

```
 52   NNVDFVRTY DTVEFLSADT KTTTVNVESK DNGKKTEVKI GAKTSVIKEK
101   DGKLVTGKDK GENGSSTDEG EGLVTAKEVI DAVNKAGWRM KTTTANGQTG
151   QADKFETVTS GTNVTFASGK GTTATVSKDD QGNITVMYDV NVGDALNVNQ
201   LQNSGWNLDS KAVAGSSGKV ISGNVSPSKG KMDETVNINA GNNIEITRNG
251   KNIDIATSMT PQFSSVSLGA GADAPTLSVD GDALNVGSKK DNKPVRITNV
301   APGVKEGDVT NVAQLKGVAQ NLNNRIDNVD GNARAGIAQA IATAGLVQAY
351   LPGKSMMAIG GGTYRGEAGY AIGYSSISDG GNWIIKGTAS GNSRGHFGAS
401   ASVGYQW*
```

FIG. 14E

```
 52   NRAASVKDV LNAGWNIKGV KPGTTASDNV DFVRTYDTVE FLSADTKTTT
101   VNVESKDNGK KTEVKIGAKT SVIKEKDGKL VTGKDKGENG SSTDEGEGLV
151   TAKEVIDAVN KAGWRMKTTT ANGQTGQADK FETVTSGTNV TFASGKGTTA
201   TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS GWNLDSKAVA GSSGKVISGN
251   VSPSKGKMDE TVNINAGNNI EITRNGKNID IATSMTPQFS SVSLGAGADA
301   PTLSVDGDAL NVGSKKDNKP VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN
351   RIDNVDGNAR AGIAQAIATA GLVQAYLPGK SMMAIGGGTY RGEAGYAIGY
401   SSISDGGNWI IKGTASGNSR GHFGASASVG YQW*
```

FIG. 14F

```
 50 SANTLKAGDNL KIKQFTYSLK KDLTDLTSVG TEKLSFSANG NKVNITSDTK
101   GLNFAKETAG TNGDTTVHLN GIGSTLTDRA ASVKDVLNAG WNIKGVKNVD
151   FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS VIKEKDGKLV
201   TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA NGQTGQADKF

251   ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA LNVNQLQNSG
301   WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE ITRNGKNIDI
351   ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV RITNVAPGVK
401   EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG LVQAYLPGKS
451   MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG HFGASASVGY
501   QW*
```

FIG. 14G